"We Help Put America Through School"

FSA Data Strategy:
Standard Student Identification Method (SSIM)

Implementation Recommendations

July 31, 2003

Meeting Objective and Agenda

Objective:

The focus of this meeting is to gain consensus on the SSIM implementation option recommendation and flush out business area impacts. The August 6th meeting will consist of breakout sessions to develop high level requirements and an overall sequencing plan.

Agenda:

- SSIM Recap
- Implementation Options Analysis
- Recommendation Consensus/Questions
- SSA Match Recommendation
- Business Area Impact Breakout
- Next Steps

Standard Student Identification Method Recap

Key Identification Problems in the Current Environment

- Unique customer records can be inappropriately merged creating privacy concerns.
- A customer's records cannot be linked appropriately preventing FSA from viewing data about a customer across all phases of the lifecycle.

Cause of the Identifier Problems

- All FSA systems may not be using the same additional identifying data. Most systems employ different rules for determining uniqueness of identities for inbound or outbound interfaces.
- Some FSA systems complete an SSN verification with SSA before data is processed; others do not perform the SSA match when new information is received.
- Changes or corrections to identifying fields (e.g., name changes) are not consistently supported across the FSA enterprise.

Standard Student Identification Method Recap

Objective

The Standard Student Identification Method seeks to establish a simple framework by which FSA and Delivery Partners can consistently identify students/borrowers, across all phases of the Student Aid Lifecycle.

High Level Requirements

- Consistently and systematically link customer records across the FSA enterprise.
- Support process changes and updates to key customer attributes (e.g., updates to SSN, First Name, Last Name, DOB.)
- Ensure student privacy protection; minimize unauthorized/unauthenticated access to student data.
- System identification requirements should not prevent valid customers from receiving aid or progressing through the repayment phase (e.g., deferments, rehabilitations, consolidations.)

Standard Student Identification Method Recap

The Standard Student Identification Method Core Team developed a ***Three-Pronged Solution:***

Leverages effective, proven identifier solutions already being used in some parts of the FSA lifecycle. Roll-out of these tools and processes consistently shall tighten controls and improve data integrity/consistency.

1. Primary Identifier Verification with the Matching Algorithm
2. Additional SSA Verification
3. Consistent Correction Processing and Error Notification

Today's session will be focused on implementation of the Matching Algorithm.

The Standard Student Identification Method Recap

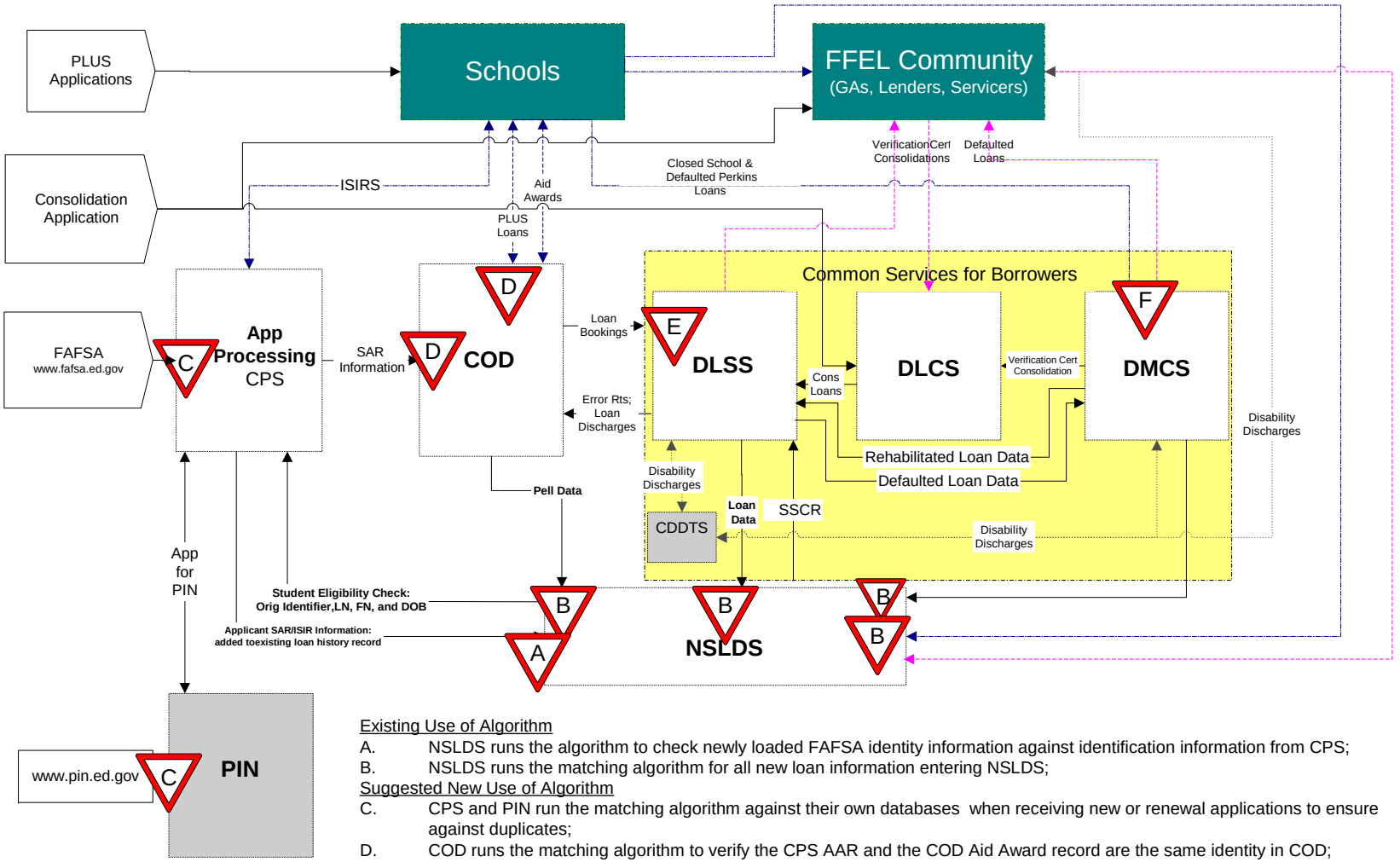
What is the "identifier" if using a matching algorithm?

- The matching algorithm requires a combination of data fields common to all systems.
- The primary identifier is the Social Security Number, but it will be verified through enterprise-wide business rules and tolerances with additional data fields: First Name, Date of Birth, and Last Name.

Why a matching algorithm?

- By employing a matching algorithm, or business rules, FSA systems can consistently identify customers throughout internal data exchange and external data acceptance.
- The use of this algorithm is a proven practice within FSA internal and external data exchange (as well as other agencies and financial institutions).
- Requires data already existing in FSA systems.
- Provides flexibility in implementation.

Recommended use of Matching Algorithm Recap



Existing Use of Algorithm

A. NSLDS runs the algorithm to check newly loaded FAFSA identity information against identification information from CPS;

B. NSLDS runs the matching algorithm for all new loan information entering NSLDS;

Suggested New Use of Algorithm

C. CPS and PIN run the matching algorithm against their own databases when receiving new or renewal applications to ensure against duplicates;

D. COD runs the matching algorithm to verify the CPS AAR and the COD Aid Award record are the same identity in COD;

E. DLSS (CSB) runs the matching algorithm to match records received from COD, DLCS, and DMCS with those existing in DLSS;

F. DMCS (CSB) runs the matching algorithm to match debts received from DLSS, Schools, or FFEL community with those existing in DMCS;

Implementation Options Analysis

Analysis Considerations

Since delivery of the High Level Design, the team has analyzed the suggested implementation options for the Matching Algorithm and the Update/Error Process, based on:

- Technical Feasibility
- Alignment with Data Strategy and other integration initiatives

Reminders

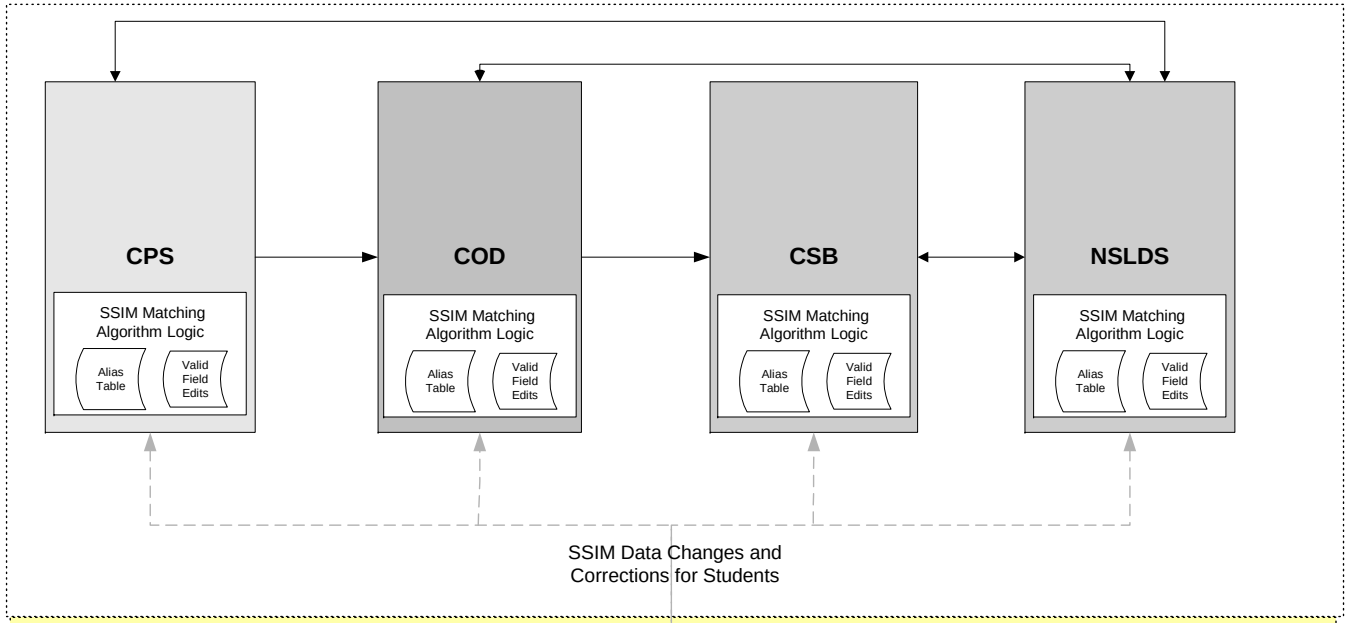
The original names of the options are listed below. These options have been changed/renamed to reflect the details of the implementation option.

- Interface Support → Application Level Implementation of SSIM Logic
- Centralized Call → Centralized Routing (e.g., EAI)
- Centralized Index → Verification with Central Data Store
- Blend (combination of options above)

Implementation Options Analysis – Overview SSIM Picture

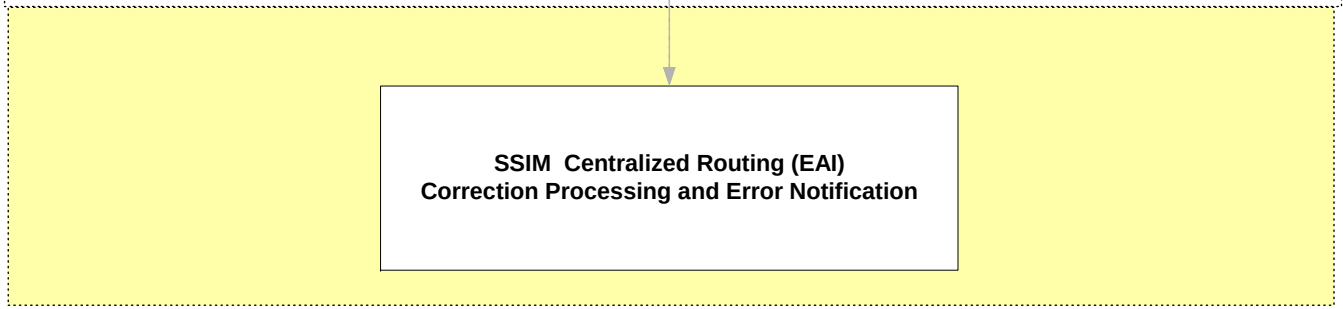
FSA Application Level

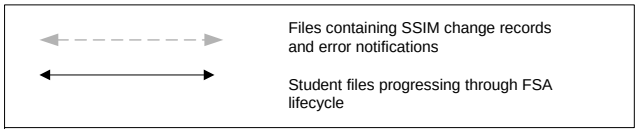
For flow of student records through the lifecycle from one FSA system to another



FSA Centralized Routing (EAI)

For changes and corrections to student SSIM data that must be communicated across the enterprise





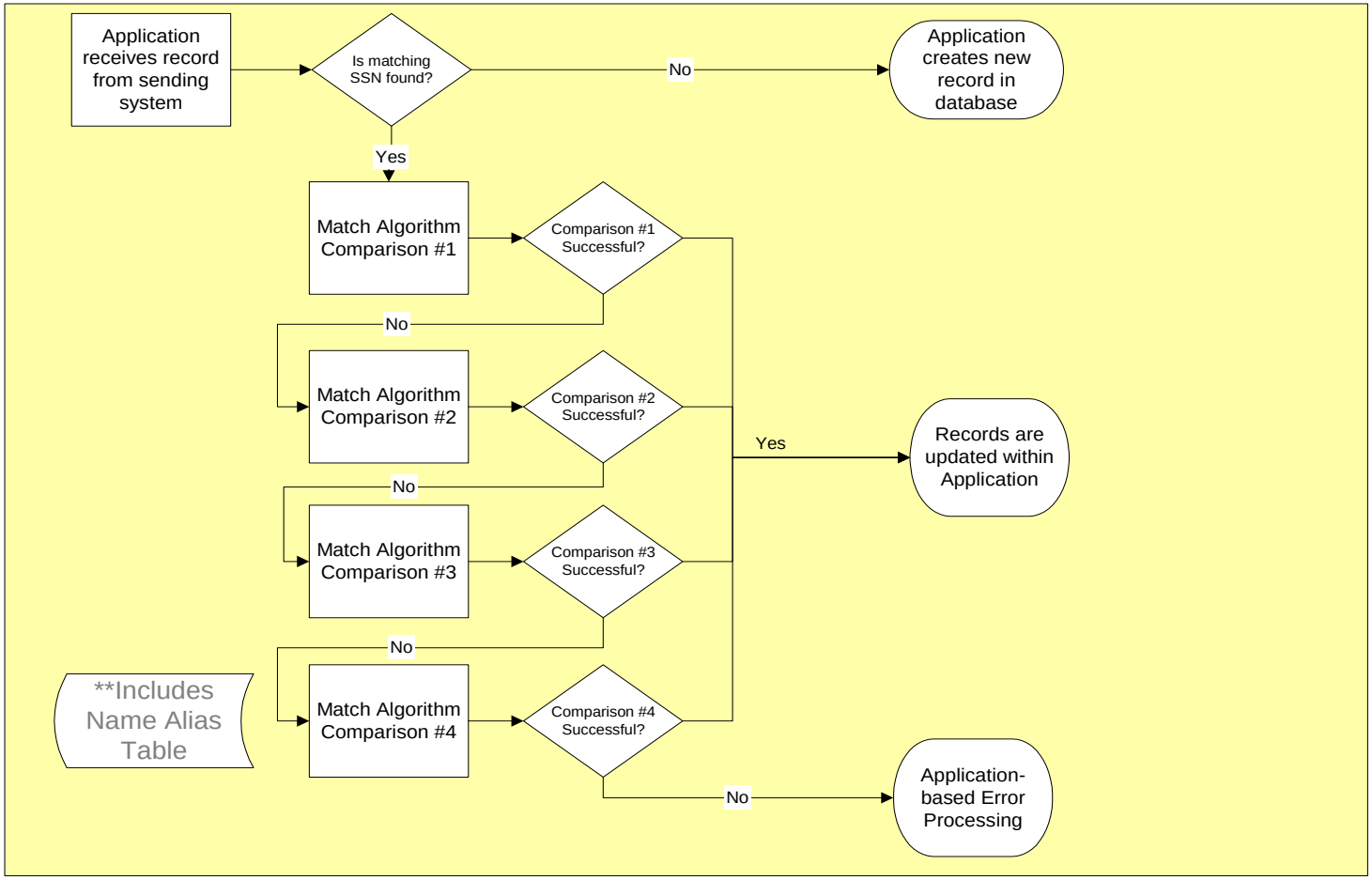
Implementation Options Analysis – Matching Algorithm

Method	Capability Details	Benefits	Drawbacks	Recommendation
Application Level Implementation of SSIM logic	Logic for the SSIM algorithm is implemented and run within each system when receiving new and updates to records.	Can be implemented in a phased approach based on the development cycles of the systems. Can begin implementation next requirements cycle.	Logic is not maintained centrally. Matching Algorithm and Alias Table would require updates in multiple locations. Difficult to manage logic and Alias Table changes centrally when implemented at the application level and maintain consistency.	X – Most efficient way to implement and realize results within the next cycle year since each system with a student data store must run the algorithm against its database when processing incoming student records.
Centralized Routing and Logic (e.g., EAI) (EAI, Service or Blend of Centralized Technologies)	Centralized logic would imbed the SSIM algorithm logic into the central layer. Each system would be required to send the relevant information from the incoming file and its own database to the central location to run the algorithm and would be responsible for processing the outcome.	Algorithm and Alias Table would be managed and maintained centrally, preventing redundant processing and allowing for easier maintenance and control of consistency. Potentially beginning the implementation of some components of the overall Data Strategy vision.	Systems would have to send data to the central location to be processed centrally and receive/process the outcome. Every system must be set up on the central architecture. Central logic would need to be able to support the volume and speed required by the systems for processing.	

Implementation Options Analysis – Matching Algorithm

This diagram depicts the use of the algorithm at the application level, for input files from one FSA system to another.

FSA Application Level

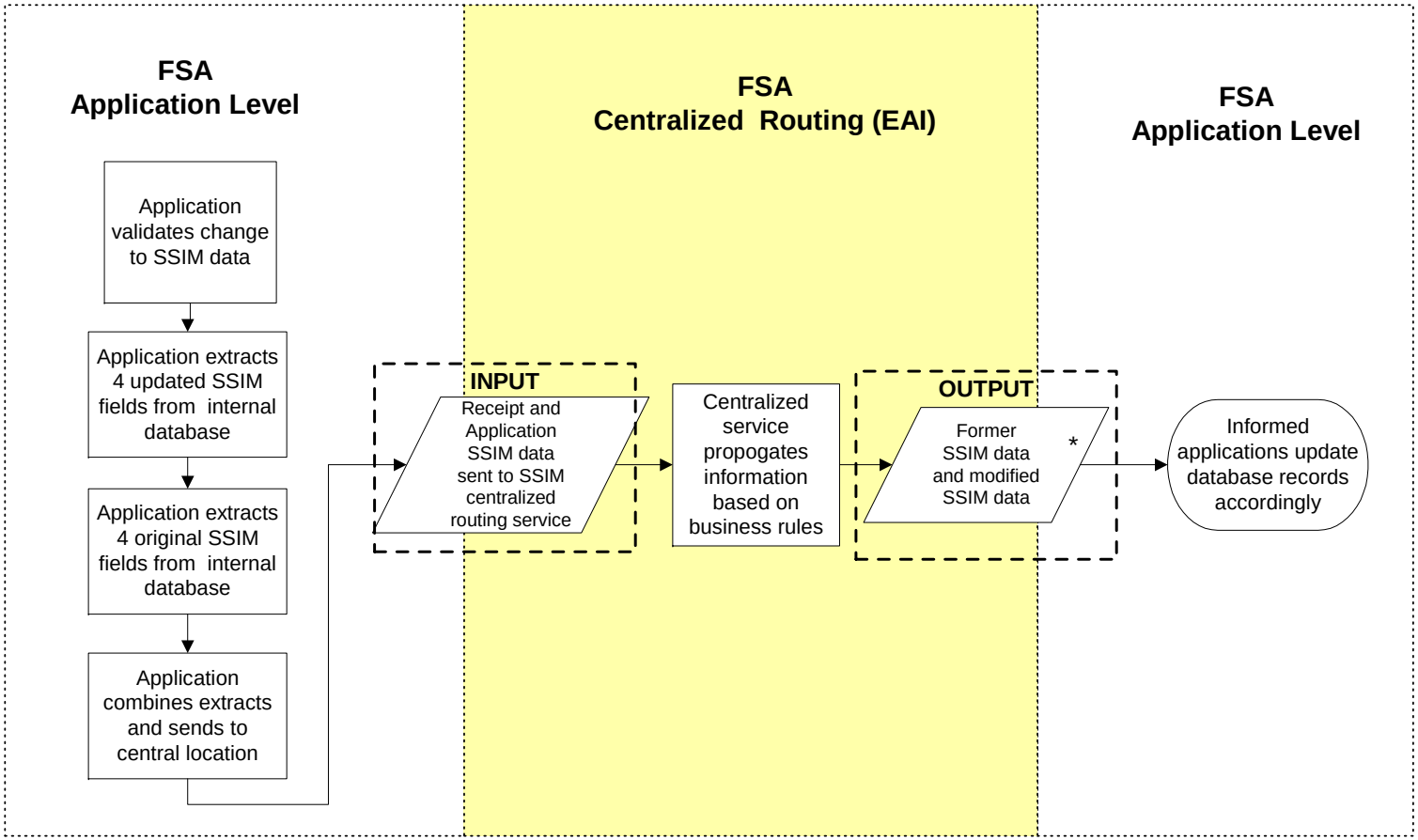


Implementation Options Analysis – Correction Processing and Error Notification

Method	Capability Details	Benefits	Drawbacks	Recommendation
Application Level Implementation of Update/Error Process	Logic for the Update/Error process is implemented and run within each system when receiving new and updates to records.	Can be implemented in a phased approach based on the development cycles of the systems.	Logic is not maintained centrally. Correction and Error Notifications need to propagate to multiple systems, making a linear application solution not the most efficient.	
Centralized Routing (EAI, Service or Blend of Centralized Technologies)	Centralized Routing (e.g., EAI) would imbed the Correction Processing and Error Notification into the central layer. Each system would be required to send updates and error notifications to the central location to process and propagate to other systems based on the predefined business rules.	Correction Processing and Error Notification would be managed and maintained centrally allowing for easier maintenance and control of consistency. Efficient solution to propagate updates/error notifications to multiple systems.	Every system must be set up on the central architecture. Central logic would need to be able to support the volume and speed required by the systems for processing.	X – Most efficient way to propagate update records and error notifications to multiple systems.

Implementation Options Analysis – Correction Processing and Error Notification

This diagram depicts the use of centralized routing (EAI) to propagate Correction Processing and Error Notification to multiple FSA systems simultaneously.



*Error Notification will be processed similarly based on business rules.

Recommendation Consensus

The recommended SSIM implementation option consists of two stages, that will allow early realization of the SSIM benefits, but also maintain alignment with the FSA overall vision.

Stage One –

A blend of the application implementation and centralized options

- Individual application's implementation of the matching algorithm option for processing input files from one system to another.
- Correction processing and error notification would be implemented through centralized routing (EAI) to allow communication/propagation to all systems.
- Implementation would begin in the next cycle year

Recommendation Consensus

Stage Two –

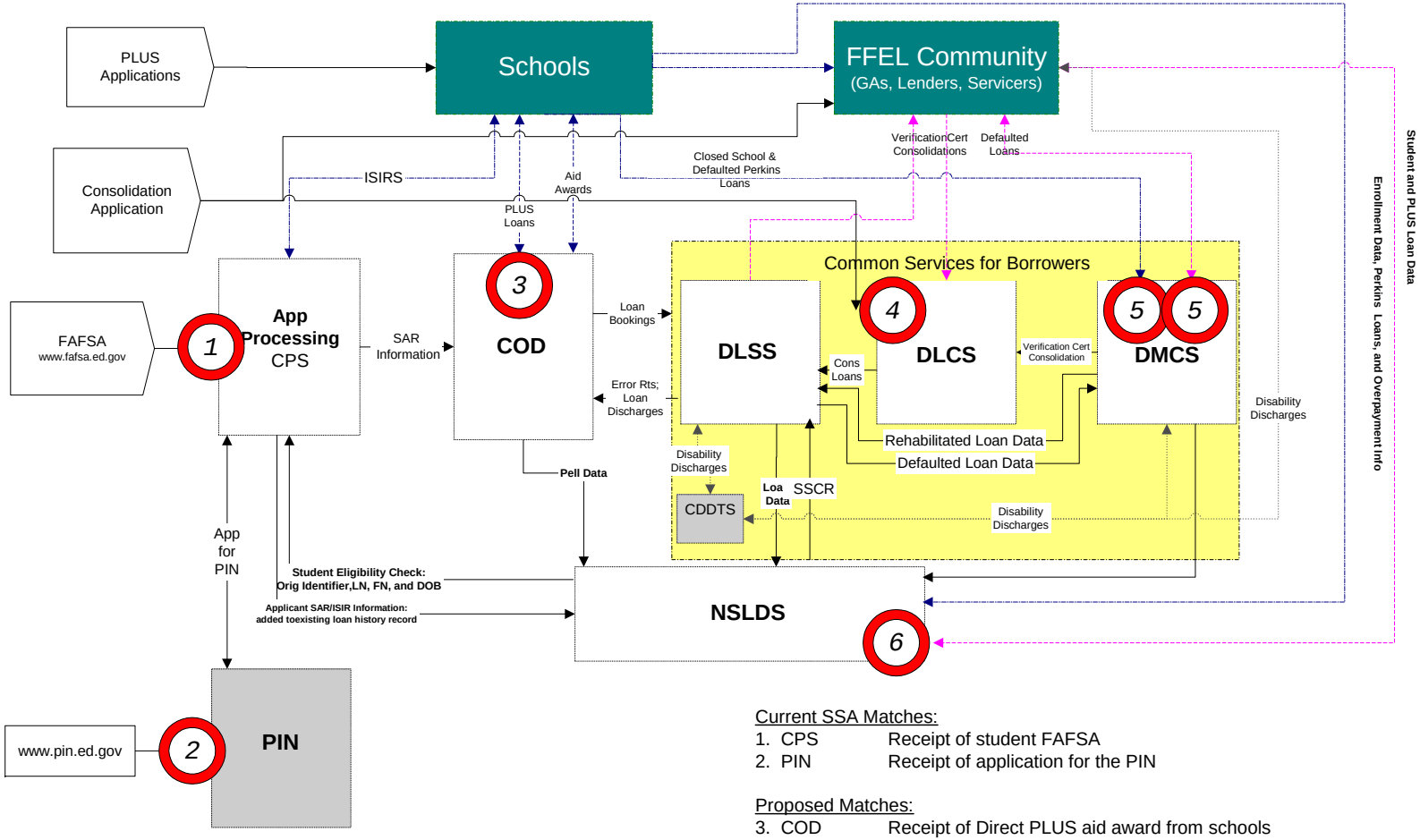
The team will create a picture and high level plan on how to include SSIM in the overall Data Strategy 3-5 year vision. This plan may include the following:

- Maintain current service for identifier changes and updates.
- Add centralized logic as a service for running the matching algorithm against the central data store (may or may not remove logic from individual applications for verification against).
- Permits validation against sending and receiving systems as well as potential central data store.
- Implementation would begin in the next 3-5 years.

Questions/Feedback?

Recommended use of SSA Identification Match

When new, unverified identities enter the FSA systems, it is appropriate to compare the record with the Social Security Administration.



Current SSA Matches:

1. CPS — Receipt of student FAFSA
2. PIN — Receipt of application for the PIN

Proposed Matches:

3. COD — Receipt of Direct PLUS aid award from schools
4. DLCS — Receipt of the Consolidation Application
5. DMCS — Receipt of new debt being loaded into the system
6. NSLDS — Receipt of FFEL PLUS records from the FFEL community

SSA Match Recommendation

For Discussion –

- The SSIM team's requirements related to additional SSA matches will be dependent on FSA policy in the coming cycle years.
- The SSIM team will seek guidance from FSA regarding PLUS applications requiring FAFSAs, or a similar central application process.
- Some policies may eliminate the need for additional SSA matching.

Breakout Sessions

For Discussion –

- The purpose of the breakout sessions are to start discussions on how the SSIM Implementation Recommendation will impact each system/group of systems.
- Please use the system impact list as a starting point (requirements and impacts incorporated from discussion in .doc attachment)

Next Steps

Next meeting is August 6:

- Exceptions and Change Processing
- High Level System Requirements/Impacts
- Sequencing Plan

Continue to collaborate with the overall Data Strategy teams and FSA BIG regarding To Be vision

Update at Software Developers Conference: August 14

Matching Algorithm Rules

The matching algorithm will be a series of 4 comparisons of identifying data. Any one successful comparison constitutes a successful match.

Comparison	SSN	First Name	Date of Birth	Last Name
1st SSN, First Name, and DOB	Current SSNs must match exactly on all 9 digits of the SSN on the student record.	3 of the first 4 significant characters of the first name must match in sequence* (in current or history), *or* alias matches exactly. Names of 3 characters or less must match exactly.	Year matches exactly; *or* Year matches plus or minus one, with month matching exactly; *or* Year matches plus or minus ten, with month and day matching exactly; *or* Date is an acceptable plug date	***N/A***
2nd Transposed First and Last Names	Current SSNs must match exactly on all 9 digits of the SSN on the student record.	Three of the first four significant characters of *last name on incoming record* must match in sequence (in current or history), the first name on the receiving record. *or* alias matches exactly. Names of 3 characters or less must match exactly.	Year matches exactly; *or* Year matches plus or minus one, with month matching exactly; *or* Year matches plus or minus ten, with month and day matching exactly; *or* Date is an acceptable plug date	***N/A***
3rd First Initial Provided for First Name w/ exact DOB	Current SSNs must match exactly on all 9 digits of the SSN on the student record.	First name begins with same letter as first initial (a name that is an initial only or an initial followed by a period, not a comma).	***Day, Month, and Year Match Exactly***	***N/A***
4th First Initial Provided for First Name w/ check on Last Name	Current SSNs must match exactly on all 9 digits of the SSN on the student record.	First character of first name matches first character of first name or first initial (current or history).	Year matches exactly; *or* Year matches plus or minus one, with month matching exactly; *or* Year matches plus or minus ten, with month and day matching exactly; *or* Date is an acceptable plug date	Five of first seven significant characters of last name match in sequence (current or history). If fewer than five characters, all characters must match.